

## КЛАСТЕРИЗАЦИЯ ИЗОБРАЖЕНИЙ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТЬЮ

Канаева И.А.

Научный руководитель: к.т.н., доцент Стоянов А.К.

Национальный исследовательский Томский политехнический университет  
[iap15@tpu.ru](mailto:iap15@tpu.ru)

### Введение

Задача кластеризации – это задача размещения входных векторов (образов) по категориям (кластерам) так, чтобы близкие векторы (схожие образы) оказались в одной категории.

Все методы кластеризации изображений условно можно разделить на две группы: статистические методы кластеризации и методы кластеризации, основанные на выделении перепадов яркости [1].

Кластеризация изображений с помощью рекуррентной нейронной сети проводится относительно яркостной компоненты изображений.

### Построение инструмента решения

Для решения поставленной задачи рассмотрим однослойную рекуррентную нейронную сеть с линейно-кусочной функцией активации вида

$$f(x) = \begin{cases} x \leq 0, & f(x) = 0; \\ x > 0, & f(x) = x. \end{cases} \quad (1)$$

Такая сеть моделирует  $n$ -слойный перцептрон, у которого одна и та же матрица весов синапсов  $\|W_{ij}^{(n)}\|$  ( $n=1,2, \dots, i, j$  – номера нейронов в соседних слоях) для всех слоёв.

При наложении условия  $\sum W_{ij}^2 < 1$  на матрицу сеть реализует сжимающее аффинное отображение многомерного входного вектора  $X$ :

$$X_n = WX_{n-1} + B \quad (2)$$

Известно, что у сжимающего отображения полного метрического пространства в себя для любой точки  $x \in X$  последовательность  $x, f(x), f^2(x) = f(f(x)), f^3(x), \dots$  сходится к единственной неподвижной точке [2]. Т.е. данная рекуррентная сеть в ходе нескольких итераций приводит любые многомерные входные значения к одной и той же устойчивой точке.

Кластеризация основывается на гипотезе компактности, то есть на предположении о том, что значения, сходящиеся к устойчивой точке за одно и то же число итераций, образуют один кластер. Оценка количества итераций, нужного для достижения устойчивой точки, производится на основе вычисления евклидова расстояния между вектором  $Y$  устойчивой точки и вектором  $X$  яркости точки изображения. Это расстояние для каждого  $L$ -мерного входного вектора  $X$ , который соответствует изображению размерами  $L=M \times N$ , определяется по формуле:

$$\rho(x, y) = (\sum_{i=1}^L (\mu x_i - y_i)^2)^{1/2}, \quad (3)$$

где  $x_i$  и  $y_i$  –  $i$ -ые координаты точек  $x$  и  $y$  в  $L$ -мерном пространстве, а  $\mu$  – масштабный коэффициент. В рамках заданной точности вычислений точки, достигающие устойчивого значения за одно и то же целое число итераций, имеют компактно расположенные евклидовы расстояния  $\rho^{(j)}$  на каком-то одном непрерывном диапазоне значений. Поэтому для кластеризации вычисляются границы таких диапазонов по формуле

$$\rho^{(1)} = \varepsilon/K, \rho^{(2)} = \rho^{(1)}/K, \dots, \rho^{(n)} = \rho^{(n-1)}/K, \quad (4)$$

где  $\varepsilon$  – точность приближения к значению устойчивой точки,  $n$  – номер диапазона, отсчитываемый от устойчивой точки,  $K$  – коэффициент сжатия отображения.

После перебора всех входных векторов изображений формируются кластеры из изображений, попавших в один и тот же диапазон расстояния  $\rho$ .

Для определения качества кластеризации рассчитывается энтропия разбиения:

$$H = \sum_i \frac{r_i}{R} \ln \left( \frac{r_i}{R} \right), \quad (5)$$

где  $r_i$  – число изображений, попавших в один и тот же  $i$ -ый диапазон, а  $R$  – общее число изображений в коллекции.

Изменение масштаба изображений приводит к новому распределению их в выделенных диапазонах и изменению энтропии.

Таким образом, определение параметров кластеризации после  $k$ -разбиений с различными масштабами  $\mu$  ( $\mu = 0.8, 0.85, \dots, 6$ ) может основываться на максимальной полученной энтропии  $\max_k H$  или на ее максимальном перепаде

$$\max_k \frac{\Delta H}{\Delta \mu}.$$

### Результаты экспериментов

Объектом изучения являлись изображения размером 50 x 50 пикселей лиц людей [3]. Исходные полноцветные изображения переводились в градации серого и использовались их представления в виде многомерных векторов значений яркости.

На рисунке 1 представлены образцы входных изображений лиц людей исходных кластеров.

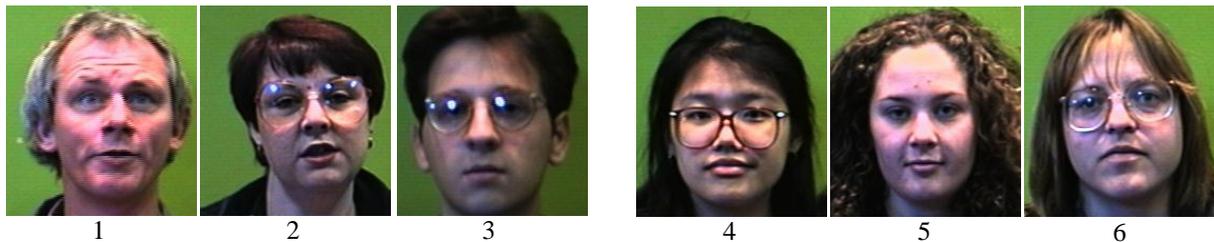


Рис. 1. Образцы изображений лиц людей

Каждое лицо представлено двадцатью изображениями, полученными при разных ракурсах, условиях освещения и отображающими разную мимику лица.

При  $K = 0.92586$  было определено 229 диапазонов, и по оценке максимальной энтропии

исходная выборка из 5 классов по 20 изображений лучше всего распределялась при масштабе  $\mu = 4.9$ . При этом результат кластеризации представлен в таблице 1.

Таблица 1. Результат кластеризации по оценке

максимальной энтропии

Кластеры эксперимента	1		2	3	4	5
Кол-во элементов в кластере	20		20	20	20	20
Полученные кластеры	1	2	3	4	5	6
	10	10	20	20	20	20

Таким образом, 4 кластера сформированы верно, а первый поровну разделился на два кластера. Однако при добавлении в коллекцию двадцати изображений 6-го человека (рис. 1) они полностью добавляются в третий выделенный кластер, содержащий изображения второго типа.

#### Вывод

Основным преимуществом данного метода является то, что для работы нейронной сети не требуется ее обучение.

Полученные результаты свидетельствуют о принципиальной работоспособности рассмотренного подхода к кластеризации изображений, но требуют дополнительных исследований, направленных на увеличение точности кластеризации.

#### Список использованных источников

1. Башков Е.А., Вовк О.Л. Кластеризация изображений методом дендрограмм / ДонНТУ, 2003. – 10 с.
2. Данилов В.И. Лекции о неподвижных точках / Российская экономическая школа, Москва, 2006 г. - 30 с.
3. Стоянов, Александр Кириллович. Применение рекуррентной нейронной сети для решения задачи кластеризации [Электронный ресурс] / А. К. Стоянов // Известия Томского политехнического университета [Известия

- ТПУ] / Томский политехнический университет (ТПУ). — 2009. — Т. 315, № 5: Управление, вычислительная техника и информатика. — [С. 144-149]. — Заглавие с титульного листа. — Электронная версия печатной публикации. — Свободный доступ из сети Интернет. — Adobe Reader. Режим доступа: [http://www.lib.tpu.ru/fulltext/v/Bulletin\\_TPU/2009/v315/i5/29.pdf](http://www.lib.tpu.ru/fulltext/v/Bulletin_TPU/2009/v315/i5/29.pdf)
4. Libor Spacek. Collection of Facial Images: Faces94 / [Электронный ресурс] / Computer Vision Science Research Projects, 2007. Режим доступа: <http://cswww.essex.ac.uk/mv/allfaces/faces94.html>, дата доступа: 20.09.2017 г.